

WEDNESDAY, AUGUST 3

finding/fighting stories, project 3 first draft critiques, debugging revisited, annotating in matplotlib, probably more things

PULL REQUESTS

<https://github.com/jsoma/playfair-projects/pulls>

Project 2, make 'em green.

FINDING YOUR STORY

(and fighting your data)

DATA FIRST VS. STORY FIRST

DATA FIRST



kromreig commented 9 days ago • edited



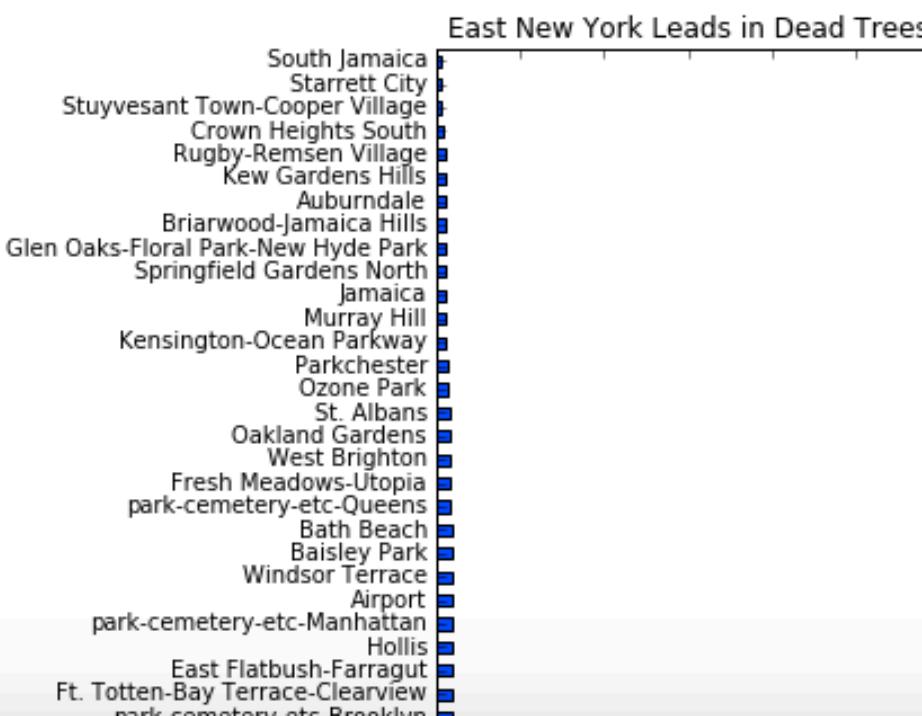
I want to talk about how many dead trees there are in East New York. The median number of dead trees in a neighborhood is 43, while East New York has nearly 300, almost double the neighborhood with the next greatest amount of dead trees.

However, I'd like to look at a few other parameters before going forward with a "EAST NEW YORK HAS SO MANY DEAD TREES" story, such as maybe tree species diversity, recent history of disease, perhaps compare the 2015 tree census data to the 1995 data and see where things stand.

Data from: <https://data.cityofnewyork.us/Environment/2015-Street-Tree-Census-Tree-Data/uvpi-gqnh>

dead trees median: 43.0

dead trees mean: 53.7616580311



DJL: Shoppers Clicking Behavior regarding eShopping Banner #47

! Open

djlee0202 opened this issue 9 days ago · 1 comment



djlee0202 commented 9 days ago • edited

+ 😊 🖊

Interesting facts of eShopping Banners

Banner Ads are big opportunities to interact with online shoppers in their path to purchase.

Therefore, understanding the efficiency of banners is critical for online retailers as they attracts different amount of visitors. I am going to study website Banner Ads.

Sample Website Banner Composition

The website is providing 12 banners at Front Page, Search Page, and Category Page.

Label

Pro

Type

Miles

No m

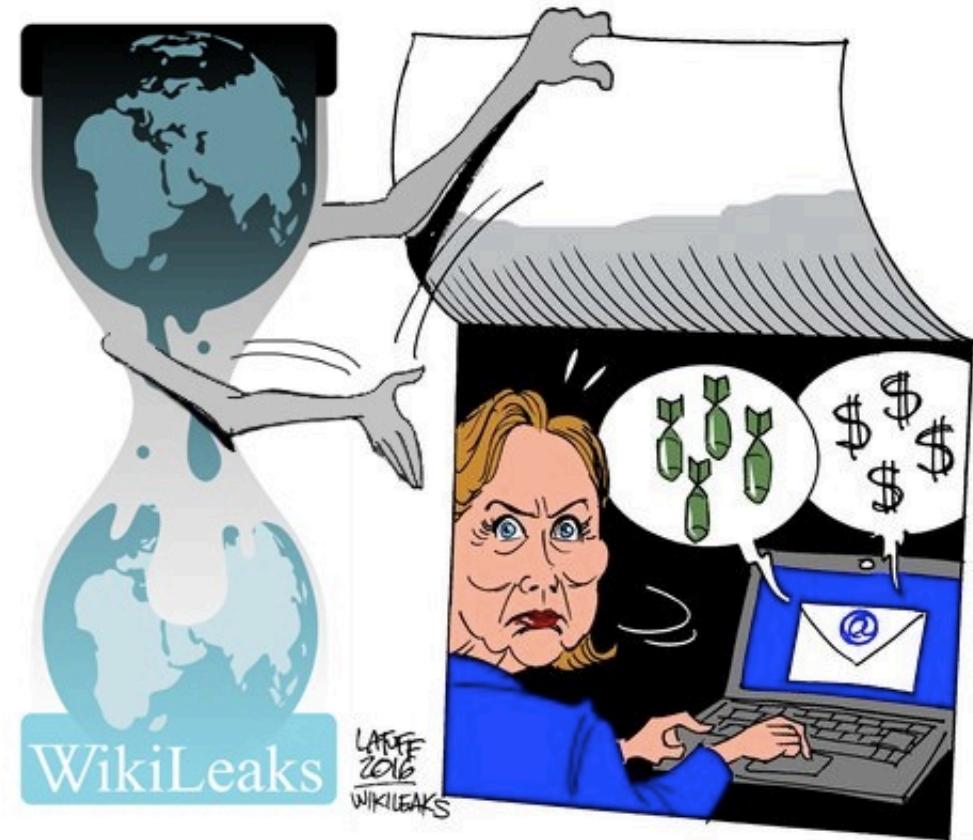
Assig

No o

3 par

Hillary Clinton Email Archive

On March 16, 2016 WikiLeaks launched a searchable archive for 30,322 emails & email attachments sent to and from Hillary Clinton's private email server while she was Secretary of State. The 50,547 pages of documents span from 30 June 2010 to 12 August 2014. 7,570 of the documents were sent by Hillary Clinton. The emails were made available in the form of thousands of PDFs by the US State Department as a result of a Freedom of Information Act request. The final PDFs were made available on February 29, 2016.



[Search by Terms in Email](#)

[Search by Email-ID](#)

Search terms throughout whole of email

You can use boolean operators to search emails.
(Example: *syria libya* will show results containing both words. *syria | libya* will show results with either words, while *syria !libya* will show results containing "syria" and not "libya".) To search an exact phrase use quotes around it like in "sensitive source". [Search techniques](#)

[Advanced Search](#)

QUARTZ GUIDE TO BAD DATA

<https://github.com/Quartz/bad-data-guide>

STORY FIRST

Ri

Billboard.com

Billboard Hot 100



Why did Billboard.com close its developer platform?



Effective immediately, Billboard will no longer offering a public API service for our charts data. New accounts will no longer be accepted, and access to existing accounts will be terminated as of May 1. 2013.

3

If you have any questions re [guoguo12 / billboard-charts](#)
feedback@billboard.com

Thank you for your interest

[Watch](#) 8 [Star](#)

[Code](#)

[Issues 0](#)

[Pull requests 0](#)

[Wiki](#)

[Pulse](#)

[Graphs](#)

- the Billboard Developer Netw

made today's tastes i
at once and it only go

Billboard has an arch
unofficial API I haven'
something with liner i
the rise of a genre as

Unofficial Python API for accessing music charts from Billboard.com. <https://pypi.python.org/pypi/billboard.py>

95 commits

2 branches

11 releases

5 co

Branch: master

New pull request

Create new file

Upload files

Find file

guoguo12 Bump to 3.0.0

Latest comm

examples/top-singles-playlist

Add example

tests

Add test for quantization

.gitignore

Initial commit

.travis.yml

Remove Python 3.5 from Travis, specify dep version

LICENSE

Initial commit

[Code](#)[Issues 70](#)[Pull requests 10](#)[Wiki](#)[Pulse](#)[Graphs](#)[Settings](#)

Closer Look at Lacrosse Legend Bill Tierney's Coaching Career #106

[Open](#)

skkandrach opened this issue 2 days ago · 2 comments

Im still working on collecting my data set..

Tierney holds the most records for National Championship titles with seven. He is also the only coach to have won a National Championship at two different institutions (i.e. Princeton and Denver). Furthermore, Tierney is the only coach to have won a National Championship with his sons sons on the roster - sharing the experience. Tierney left Princeton for Denver to spread lacrosse out west and bring more awareness to a sport that is/was dominated by the East Coast. Denver is now the only team outside of the Eastern Time Zone (and furthest team out west) to have won a National Title. I want to analyze the progression of Tierney's career, including how long it took him to win the titles after being appointed head coach, time in between titles, how many seasons he boasted an above .500 record (winning season percentage), how many appearances he's had in the National Tournament including Final Fours and Tournament Finals.

Im still working on collecting my data set..

ch

Milestone

No milestone

Assignees

No one—ass

4 participants**Notifications**

Avocado toast hotspots in New York City #77

[Edit](#)[New Issue](#)

! Open [juneeseo](#) opened this issue 3 days ago · 3 comments



juneeseo commented 3 days ago • edited



Link to the data set

[Twitter APIs](#)

Data set

I queried tweets that mention avocado toast(+ avo smash, avocado smash) from 1/1/2016 to 7/31/2016. The outcome table has xy coordinates so I can throw those in CartoDB and see where New Yorkers are getting their avocado toast.

	Unnamed: 0	lat	lng	time
0	0	40.747730	-73.986620	2016-07-28 14:15:19
1	1	40.719100	-74.000130	2016-07-28 13:29:32
2	2	40.761880	-73.968260	2016-07-28 12:22:11
3	3	40.729180	-74.000970	2016-07-26 14:19:13
4	4	40.724334	-74.002332	2016-07-24 14:26:37
5	5	40.744158	-73.988747	2016-07-23 16:56:22
6	6	40.703828	-74.011396	2016-07-23 16:22:09

Data request

I am facing an issue. It only gives me most recent 7 tweets. There is no way only 7 people mentioned avocado toast in New York City this year. Something must be wrong. I can share my code.

Possible headlines

Labels

Project 2

Request: Data

Type: Pitch

Milestone

No milestone

Assignees

No one—assign yourself

4 participants



Notifications

[Unsubscribe](#)

You're receiving notifications because you were mentioned.

[Lock conversation](#)

Avocado Toast Hot Spot in Twitter search api always return geo="<null>"

<https://twittercommunity.com/t/twitter-search-api-always-return-geo-null/66166>

Twitter Developers / Forums

Twitter Developers

Sign Up Log In

Twitter search api always return geo="<null>"

REST API

pungme May 1

Hi, I'm the builder of tweet map app (<https://itunes.apple.com/us/app/tweetmap-for-iphone/id1033296267?mt=8>)

Last week and months ago it used to return with the geo data of the tweets. I search the tweets using this code from wrapper, STTwitter

```
[_twitter getSearchTweetsWithQuery:@"" geocode:geocodeString lang:@"" locale:@"" resultType:@""] errorBlock:^(NSError *error) {
```

Since last week, all the tweets return with geo="" even i tried with my own tweet with location enable, my own tweet does not have geo data in it.

Is this the issue from twitter side? Or are there some changes in the api?

Cheers,
Pung May 4

May 1 1 / 7 May 2

May 4

created May 1 last reply May 4 6 replies 284 views 3 users 4 likes 2 links



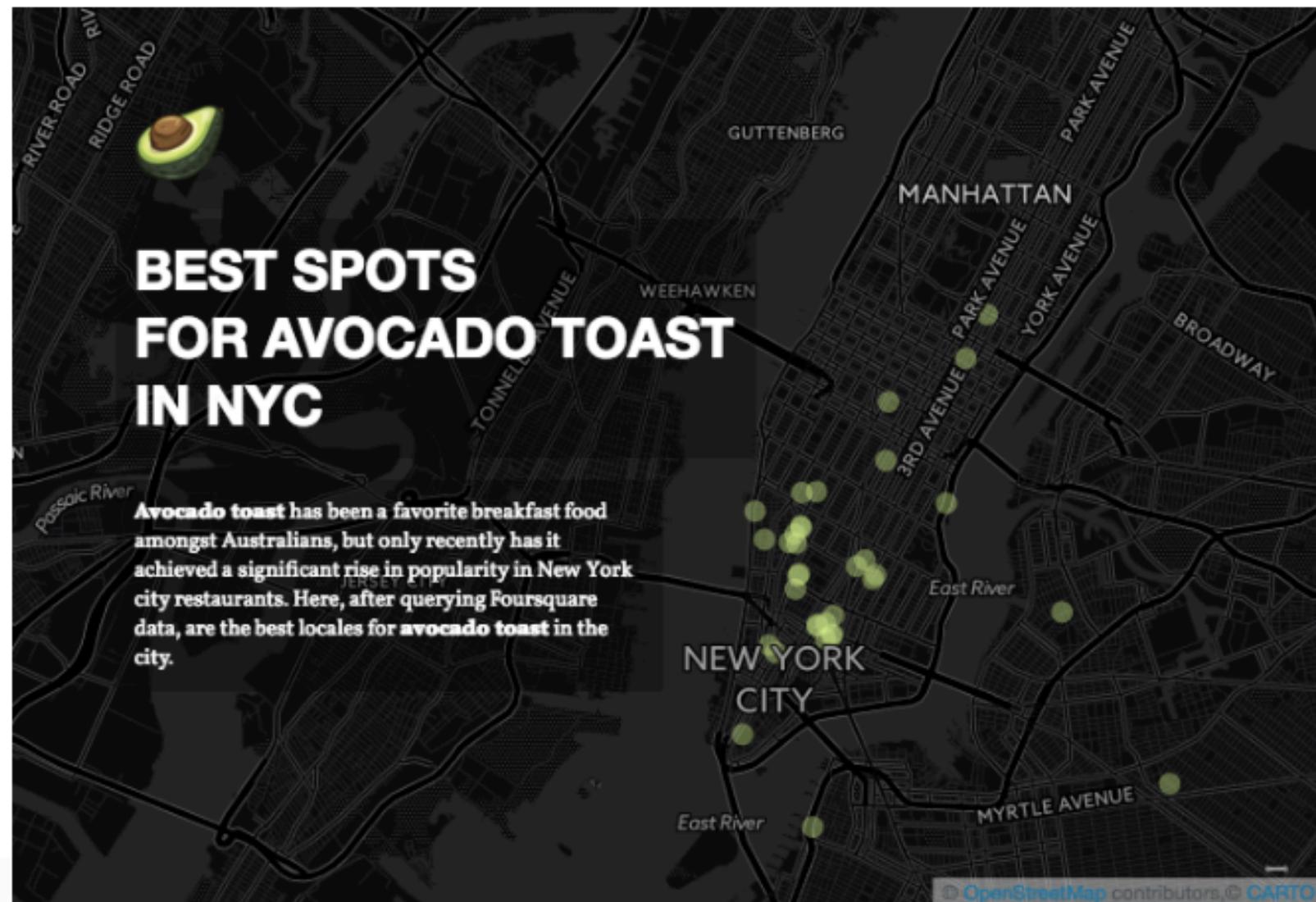
junee seo commented a day ago • edited



Map Publish

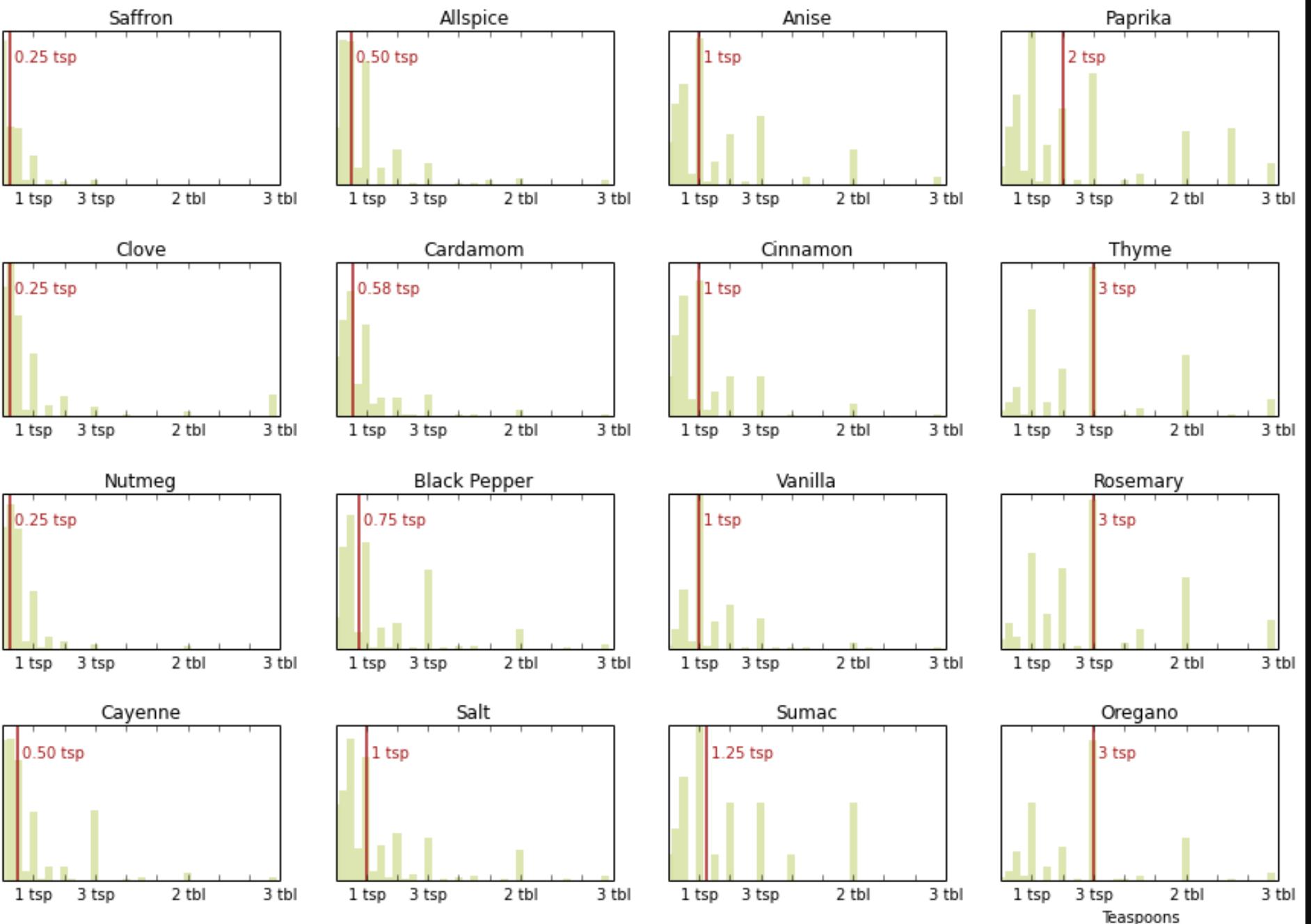
Avocado Toast Hot Spot in New York City

Screenshot



PICK A BEAT

Measurement Distribution of Spices in Recipes
@dangerscarf | omgmsg.com





tell us what's in your pantry:
sharp cheddar cheese, multigrain bread,
9 onions, pepper, dijon mustard

and we'll tell you what you can make

like macgyver, but for food

search recipes

I made this site because I was tired of having a pantry full of food but no clue what to cook with it



the tour!

Snacksby might seem a little intimidating at first, so I put together a little tour you might be interested in.

[Take the tour!](#)

Thanks, everyone! Keep the feedback coming! I know a lot of stuff is broken on the site, but I just didn't think it would get popular so quickly! Drop me an email at jonathan.soma@gmail.com if you notice something is astray or if you have a feature request (a lot of stuff isn't quite up to par) and I'll be sure to check it out.

Snacksby is a different, easier kind of recipe site. You tell us what you've got in your fridge, in your pantry, or just what you're looking to find in a recipe, even how much of each ingredient and wham!, we let you know what your options are.

Recipes and ingredients are tagged with **snacksonomies**, which are tags like **lowfat** or **vegetarian** that help you find what you might like better and faster.

[home](#)

[recipes](#)

[ingredients](#)

[snacksonomies](#)

[search](#)

[help](#)

[my account](#)

[about](#)

- ⓘ **Rise in U.S. jail and prison inmate population** Project 3 Type: Story 1
- ⓘ **NY Prison Population** Project 2 Type: Story 3
 - #100 opened 2 days ago by snajmabadi 

- ⓘ 2 Open 1 Closed Author ▾ Labels ▾ Milestones ▾ Assignee ▾ Sort ▾
- ⓘ **Accessibility of Paris' subway network** Project 3 Type: Story 1
 - #129 opened 9 hours ago by oargueso
 - ⓘ **How accessible is Madrid's subway network?** Project 2 Type: Story 4
 - #102 opened 2 days ago by oargueso 

BE SPECIFIC

[Code](#)[Issues 70](#)[Pull requests 10](#)[Wiki](#)[Pulse](#)[Graphs](#)[Settings](#)

Closer Look at Lacrosse Legend Bill Tierney's Coaching Career #106

[Open](#) skkandrach opened this issue 2 days ago · 2 comments



skkandrach commented 2 days ago

Bill Tierney is a literal legend in the college lacrosse world. Having never played lacrosse, Tierney holds the most records for National Championship titles with seven. He is also the only coach to have won a National Championship at two different institutions (i.e. Princeton and Denver). Tierney is the only coach to have won a National Championship with his sons sharing the experience. Tierney left Princeton for Denver to spread lacrosse awareness to a sport that was dominated by the East Coast. Denver is now the only team in the Eastern Time Zone (and furthest west) to have won a National Title. I will show the progression of Tierney's career, including how long it took him to win the titles after becoming head coach, time in between titles, how many seasons he boasted an above .500 record (season percentage), how many appearances he's had in the National Tournament including Tournament Finals.

I'm still working on collecting my data set..

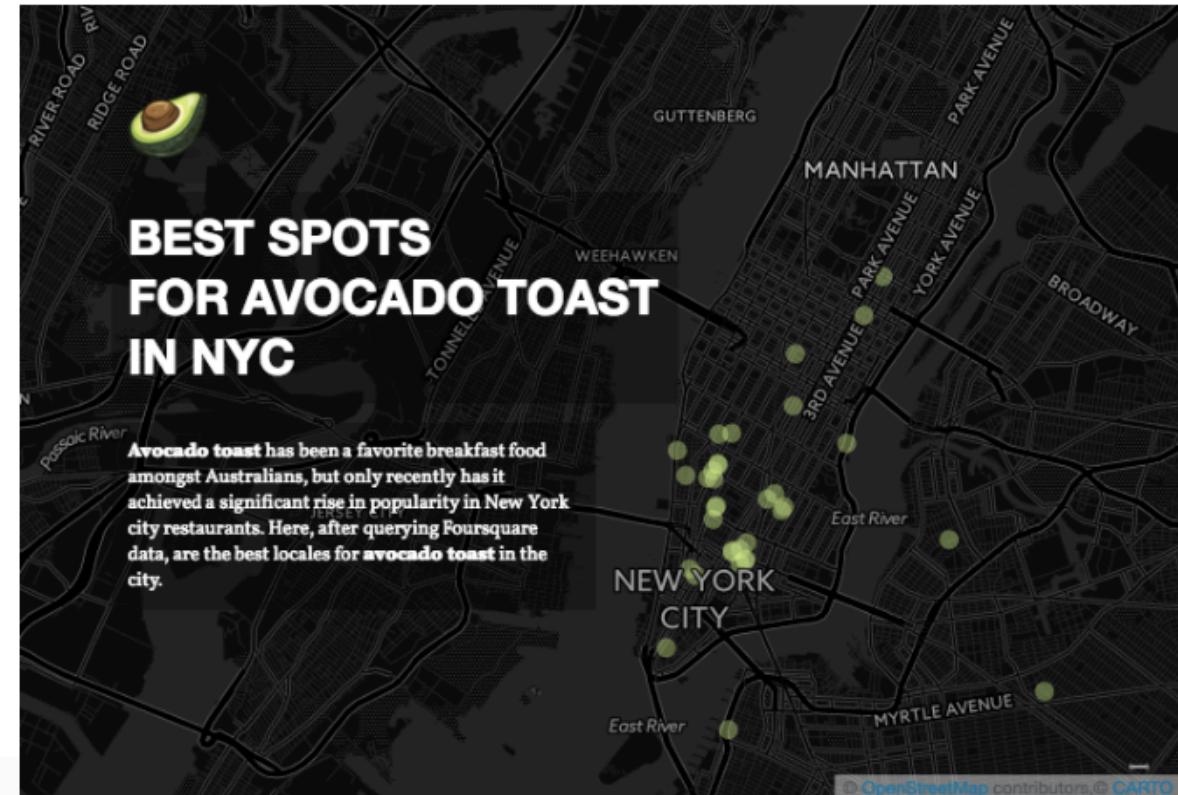


juneeseo commented a day ago • edited

Map Publish

[Avocado Toast Hot Spot in New York City](#)

Screenshot



COMBINE DATASETS



This repository Search

Pull requests Issues Gist



jsoma / playfair-projects

Unwatch 3

Star 1

Fork 26

Code

Issues 70

Pull requests 10

Wiki

Pulse

Graphs

Settings

Does a Military Coup affect Economic Growth of a Country? #48

Edit

New Issue

Open

HarshaDevulapalli opened this issue 7 days ago · 1 comment



HarshaDevulapalli commented 7 days ago



I discovered a dataset on Military Coups across the world, [thanks to Jeremy Singer Vine's Data is Plural.] - Two political science professors at the University of Kentucky are compiling a dataset of coup attempts. So far, the dataset covers both successful and unsuccessful attempts from 1950 to late 2015.

A coup as defined by this dataset is "illegal and overt attempts by the military or other elites within the state apparatus to unseat the sitting executive," and successes as episodes in which the perpetrators control power for at least 7 days. ".

During those 65+ years, coup plotters have been foiled about half the time, with 236 victories and 238 failures.

The impact of these coups on economic growth has not been studied much. We could try seeing the impact of this on a parameter such as GDP to see if they do have an impact or not.

I will annotate the GDP data I have made for my story earlier - along with this data on successful coups

Labels

Type: Pitch

Milestone

No milestone

Assignees

No one—assign yourself

3 participants



Notifications

COMPARED TO WHAT?



jowang0319 commented 2 days ago • edited



Dataset:

- [dataset](#)
- [dataset2](#)

I want do visualize something about this dataset about all coups around the world since 1950s. Maybe where did all these coups happen, and the proportion of success, etc...

Possible headlines: No idea yet..

Variable	Description
country	Country name
ccode	Country codes corresponding to Gleditsch & Ward (1999)
verified	1 = verified as an unsuccessful coup attempt; 2 = verified as a successful coup attempt; missing = not verified as a coup attempt
year	Year of coup attempt defined by primary author
month	Month of coup attempt defined by primary author; -9 = primary author did not list a month for the event
day	Day of coup attempt defined by primary author; -9 = primary author did not list a day for the event
source ²	Primary source used to document the case. "None" indicates that we were unable to find any documentation supporting the case. "Unclear" indicates that the event did not include enough information to be documented. Most of these cases came from authors including only the year for an event (missing month and/or year) in a state that had multiple coups (e.g., observations 15-16). The authors retain all original copies of source documents (available upon request).

mcgowan	Describes events coded by McGowan (2007). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
csp	Describes events coded by Marshall & Marshall (2007). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
bs	Describes events coded by Belkin & Shofer (2003). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
banks	Describes events coded by Banks (2001). 2 = event coded as a successful coup; missing = event not coded. Note: Banks (YEAR) does not code unsuccessful coup attempts.
moreno	Describes events coded by Moreno et al. (2004). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
finer	Describes events coded by Finer (1988). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
janowitz	Describes events coded by Janowitz (1977). 2 = event coded as a successful coup; missing = event not coded. Note: Janowitz (1977) does not code unsuccessful coup attempts.
kennedy	Describes events coded by Kennedy (1974). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
ferguson	Describes events coded by Ferguson (1987). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
okane	Describes events coded by O'Kane (1987). 2 = event coded as a successful coup; missing = event not coded. Note: O'Kane (1987) does not code unsuccessful coup attempts.
thompson	Describes events coded by Thompson (1973). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
luttwak	Describes events coded by Luttwak (1969). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded
lunde	Describes events coded by Lunde (1991). 1 = event coded as an unsuccessful coup; 2 = event coded as a successful coup; missing = event not coded

no_evidence	Coded 1 if no evidence was found to verify the event, 0 otherwise.
rebellion	Coded 1 if the event is better described as a rebellion than a coup, 0 otherwise. Most of these events were challenges by actors outside of the state apparatus.
no_overt	Coded 1 if there was no overt action taken to overthrow the executive, 0 otherwise. Most of these events are better described as coup plots or rumors.
legal	Coded 1 if the event was legal, 0 otherwise. Most of these events were the result of impeachment-like proceedings.
no_exec	Coded 1 if the event was not meant to overthrow the chief executive, 0 otherwise.
protest	Coded 1 if the event is better described as a protest or mass demonstration, 0 otherwise. The majority of these cases are coded 1 because the opposition did not include a member of the state apparatus.
autogolpe	Coded 1 if the sitting executive seized power for him/herself, 0 otherwise.
invasion	Coded 1 if the attempt to unseat the sitting executive came from an external force, 0 otherwise.
unclear	Coded 1 if it is unclear what event the case is referring to, 0 otherwise. Most of these cases came from authors including only the year for an event (missing month and/or year) in a state that had multiple coups (e.g., observations 28-29).

Biggest Chatter Boxes in Swiss Parliament #76

! Open

barjacks opened this issue 4 days ago · 4 comments



barjacks commented 4 days ago • edited



Summary:

This spring German speaking, mostly right wing politicians, were complaining about how long and how much their French colleagues talk in parliament. Do French speakers really spend more time talking to parliament?

Source:

After every parliamentary session - which is usually 4 a year, spring, summer, autumn and winter - The Swiss parliamentary services publishes [a PDF with approximately 1000 pages](#) containing every single discussion in parliament.

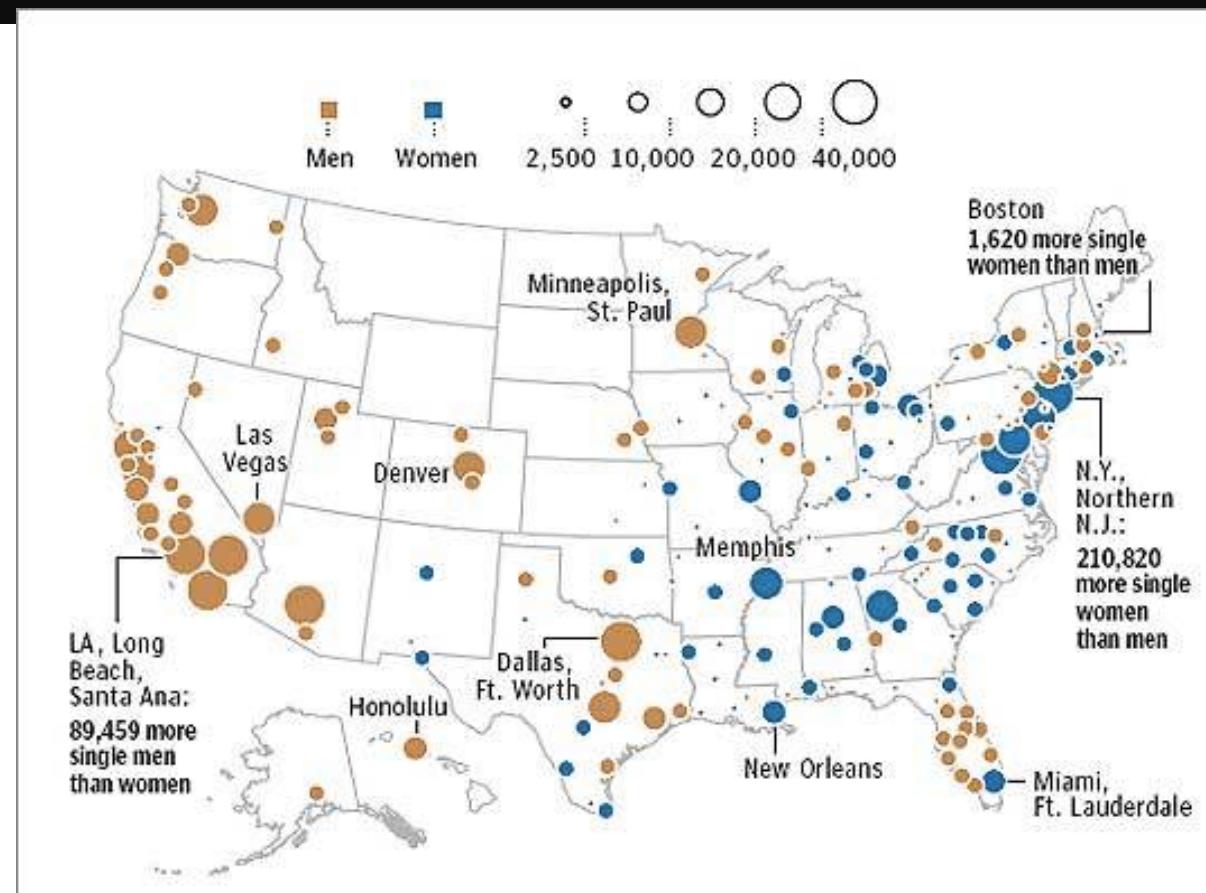
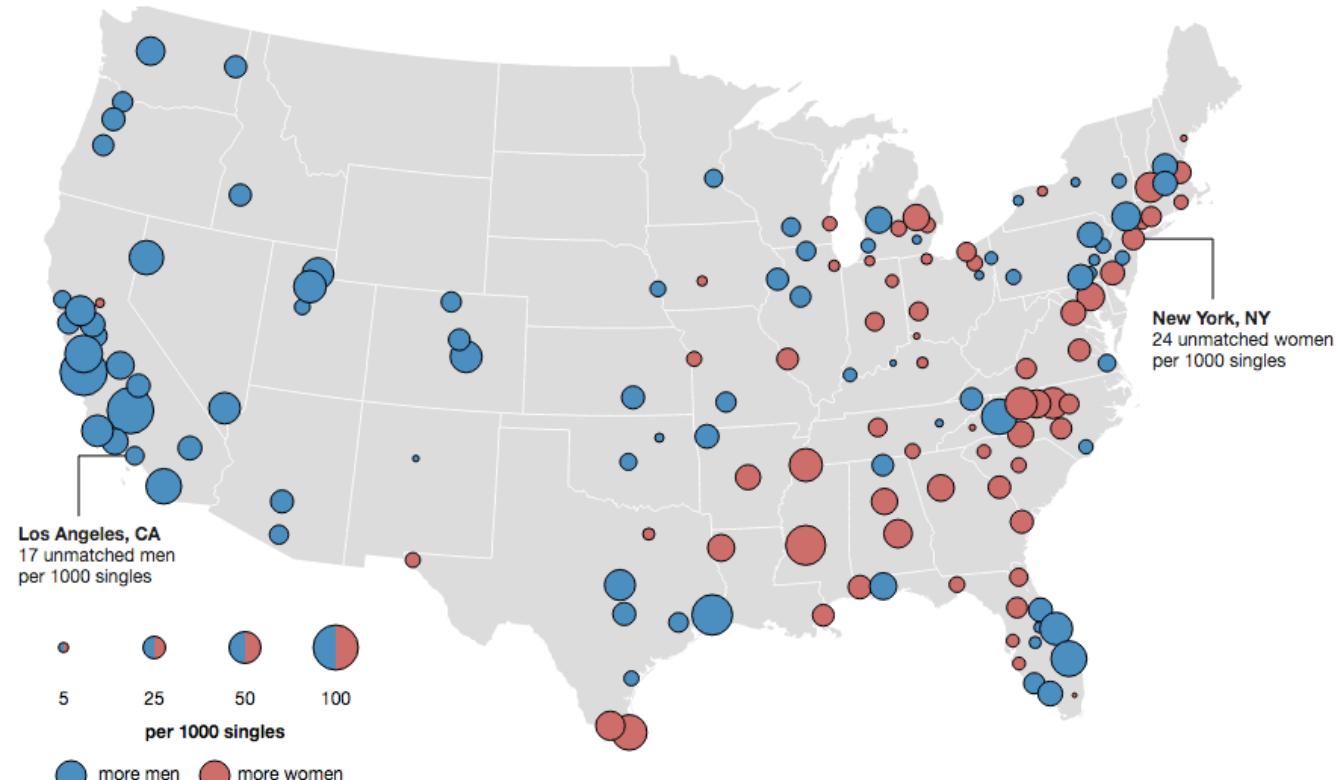
Tools and Process:

Using [PyPDF2](#) and Regular Expressions I will scrape the parliament PDF, create a dictionary with Name, Party, Canton, Length of speeches and separate one to look up the speeches themselves. The politician's I will add later on.

**STANDING ON THE
SHOULDERS OF GIANTS
(OR EVEN NORMALLY-
SIZED PEOPLE)**

Singles, 18-64

Showing cities of 250k+ people and relative values



AARON ATENCIO/GLOBE STAFF

A better NBA contracts evaluation, maybe #40

! Open

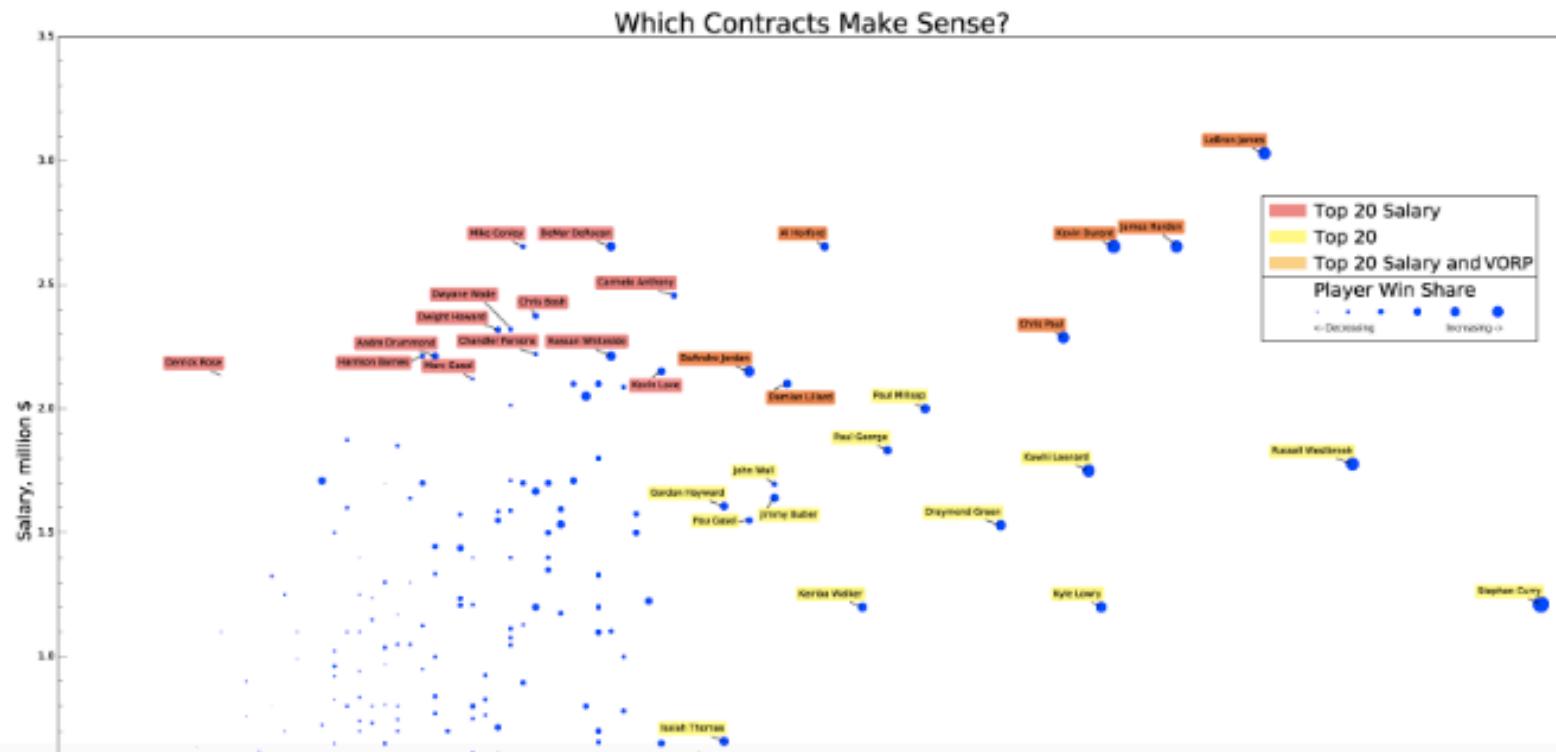
spm2164 opened this issue 9 days ago · 6 comments



spm2164 commented 9 days ago



Since points per game is barbarous and terrible, I decided to do some charts evaluating NBA contracts using better, more advanced stats. In this one I used VORP, which means value over replacement player. A replacement player is so bad that they are worth -2, and are actively making their team lose. I then highlighted the top VORP players by name and top contracts by dollar amount to see where they intersected, and where the really bad outliers were. Here it is:



Koch University Funds: Return on Investment? #96

! Open

kromreig opened this issue 2 days ago · 5 comments



kromreig commented 2 days ago • edited



I found a [dataset](#) of the total Koch brothers contributions to US Universities, from 2005 to 2014.

I want to combine it possibly with some data on research output by said universities, or median student income, or median student indebtedness, most common areas of study, etc.

I'm curious to see what the connection behind the money could be, as some of the donations are quite hefty (max is \$45 million) and others are not (min \$1,000). So the question here is, why the huge differences, what are the differentiating factors, and what are they getting in return? TANSTAAFL

Also, I found a ton of analysis on the subject, but not much data besides the original set.

[Salon's article](#)

[The Atlantic's article](#)

CHARLES KOCH FOUNDATION VS. ACADEMIC FREEDOM:

Before, and since, Greenpeace's 2014 report documenting [Charles Koch and Academic Freedom](#) issues on Koch-funded campuses, commentary has amassed from students and professors on impacted campuses, in addition to analyses from reporters, bloggers and researchers.

PROFESSORS DISCUSS KOCH-UNIVERSITY ARRANGEMENTS THAT BREAK ACADEMIC FREEDOM GUIDELINES:

- [Koch and Academic Freedom](#) - University of Illinois Professor John K. Wilson, Academe, American Association of University Professors (AAUP), 2015.
- [Stacking Influence at Taxpayer Expense](#) - Wright State University Professor Martin Kich, Ph.D, for AAUP Ohio Chapter, 2015.
- [AAUP Expresses Strong Concern Over Renaming GMU Law School and Issues of Shared Governance](#), Anita Levy, Ph.D, Associate Secretary, American Association of University Professors, 2016.
- [FSU Progress Coalition: Some Substantial Changes Have Been Made to the University's Donor Policy](#) - Wright State University Professor Martin Kich, Ph.D, for AAUP Ohio Chapter, 2015.
- [KU supports faculty members' rights](#) - Bernadette Gray-Little, Kansas University Chancellor, 2015.
- [Academic Integrity](#) - Kansas University professor Ron Barrett-Gonzales, Ph.D, for AAUP Kansas Chapter, 2014.
- [AAUP Supports Students' Request](#) - Kansas University professor Ron Barrett-Gonzales, Ph.D, 2014.
- [FSU learns from Koch association](#) - Florida State University professor Ray Bellamy and professor emeritus Kent Miller, 2014.
- [FSU should be proud of My View by students](#) - Florida State University professor emeritus Marc E. Freeman, 2014.
- [Fine Print, Restrictive Grants, and Academic Freedom](#) - Florida State University professor Ray Bellamy (M.D.) and professor emeritus Kent Miller, 2012.

STUDENTS ON CHARLES KOCH GRANTS VIOLATING ACADEMIC FREEDOM PRINCIPLES:

FLORIDA STATE UNIVERSITY (FSU):

- [Progress Coalition report uncovers alleged violation of university donor policy](#) - Joseph Zaballos, FSU News, 2015.
- [Koch-funded positions raise serious concerns](#) - Lakey, Tallahassee Democrat, 2015.
- [Reply to John Thrasher on Koch agreements and Transparency at FSU](#) - Lakey, FSU Progress Coalition, 2015.

INTERACTIVE DATABASE: Koch Foundation Funding to Colleges and Universities

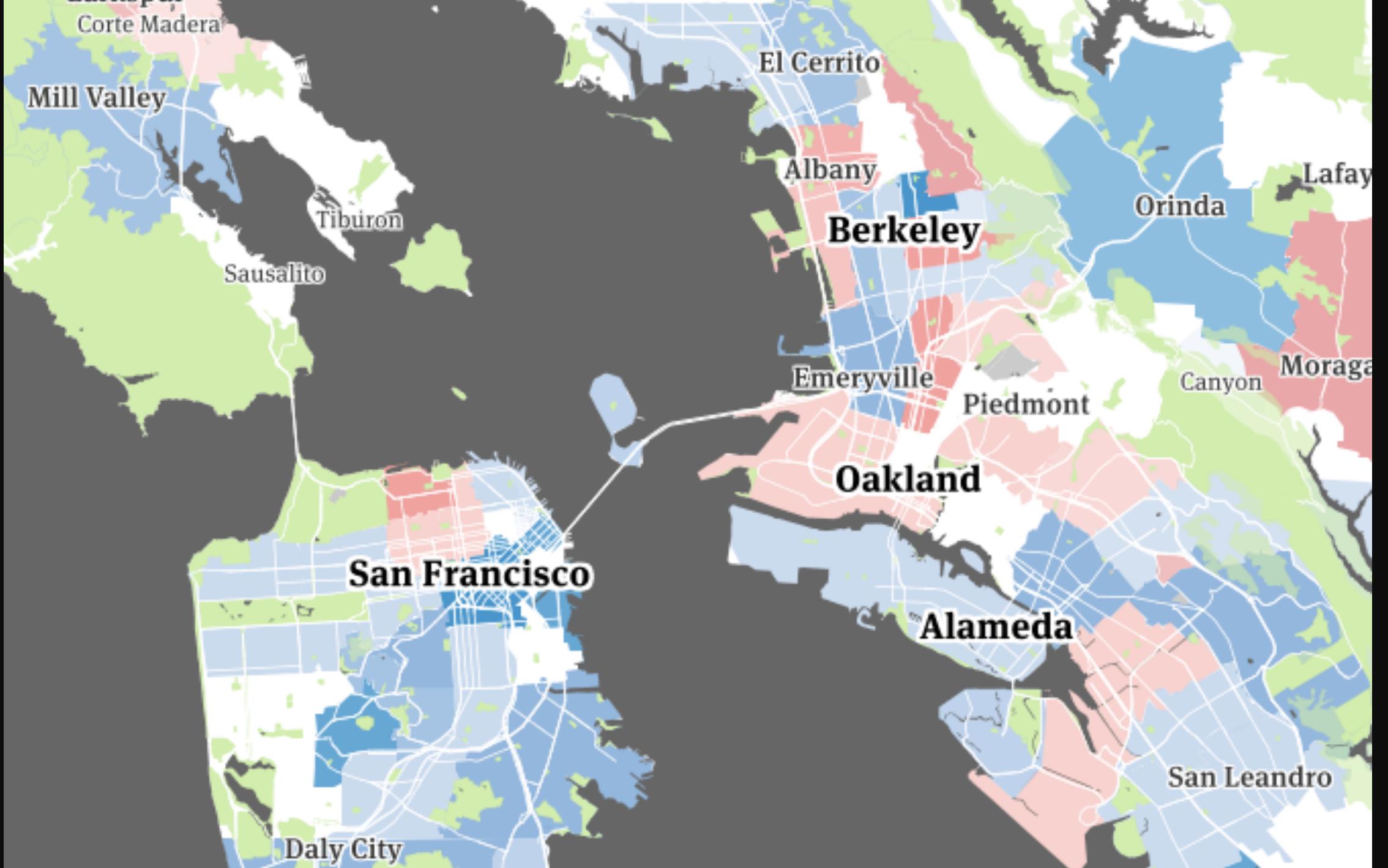
Key findings:

- The number of universities funded by Koch skyrocketed from just seven in 2005 to about 250 in 2014
- Koch compensated a Florida State University Ph.D student over \$500,000 while Koch and the FSU Economics Department negotiated a contract together
- Professors in Koch-funded departments are involved in climate science denial despite lack of scientific credentials
- Political groups hosted by Koch-funded universities are used to attack clean energy policies and pollution regulations

Lindsey Berger contributed research for this report. Authored by Connor Gibson.

Over the last several years, Greenpeace has [exposed](#) how foundations controlled by Kansas billionaires Charles and David Koch, owners of Koch Industries, have sent tens of millions of dollars to groups that deny climate change science, undermine policy solutions to the problem,

CHANGE THE SCALE



ASK SOMEONE

[Code](#)[Issues 70](#)[Pull requests 10](#)[Wiki](#)[Pulse](#)[Graphs](#)[Settings](#)

[Project 3] Causes Of Death Around The World #84

[Open](#)

honjy opened this issue 2 days ago · 3 comments



honjy commented 2 days ago • edited



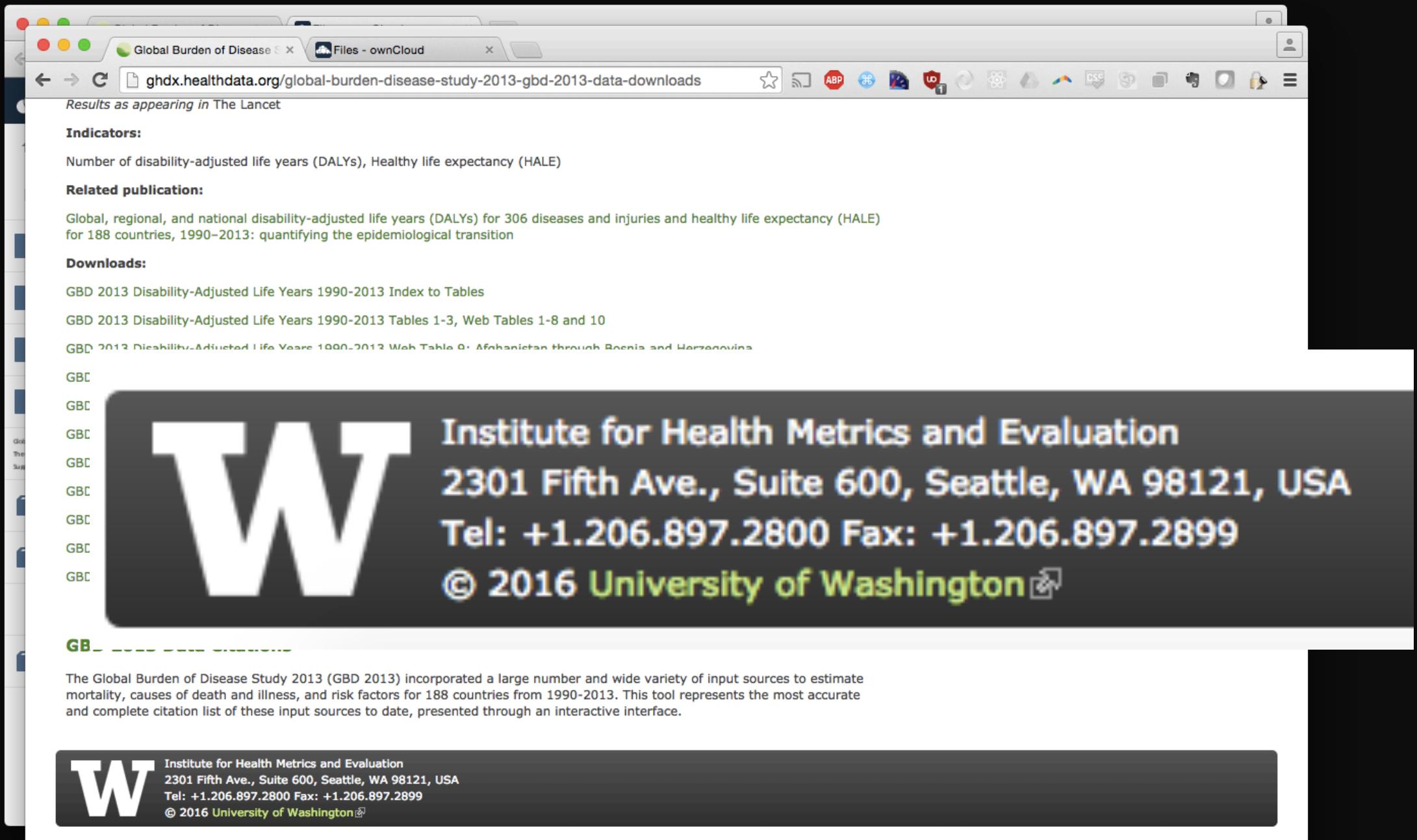
Using data from the [Global Burden of Disease Study 2013](#) I'd like to create a map showing the leading cause of death in every country in the world. Ideally I'd create several maps, showing the leading cause of death in 1990, 1995, 2000, 2005, 2010 and 2013, to see how the world has evolved in the last two decades. I'd like to learn how to create an interactive map so users can see the evolution of death patterns.

Here's an example of a map done using the same [dataset](#).

Possible headlines:

1. Leading causes of death around the world
2. This is how you'll most likely die, based on where you live

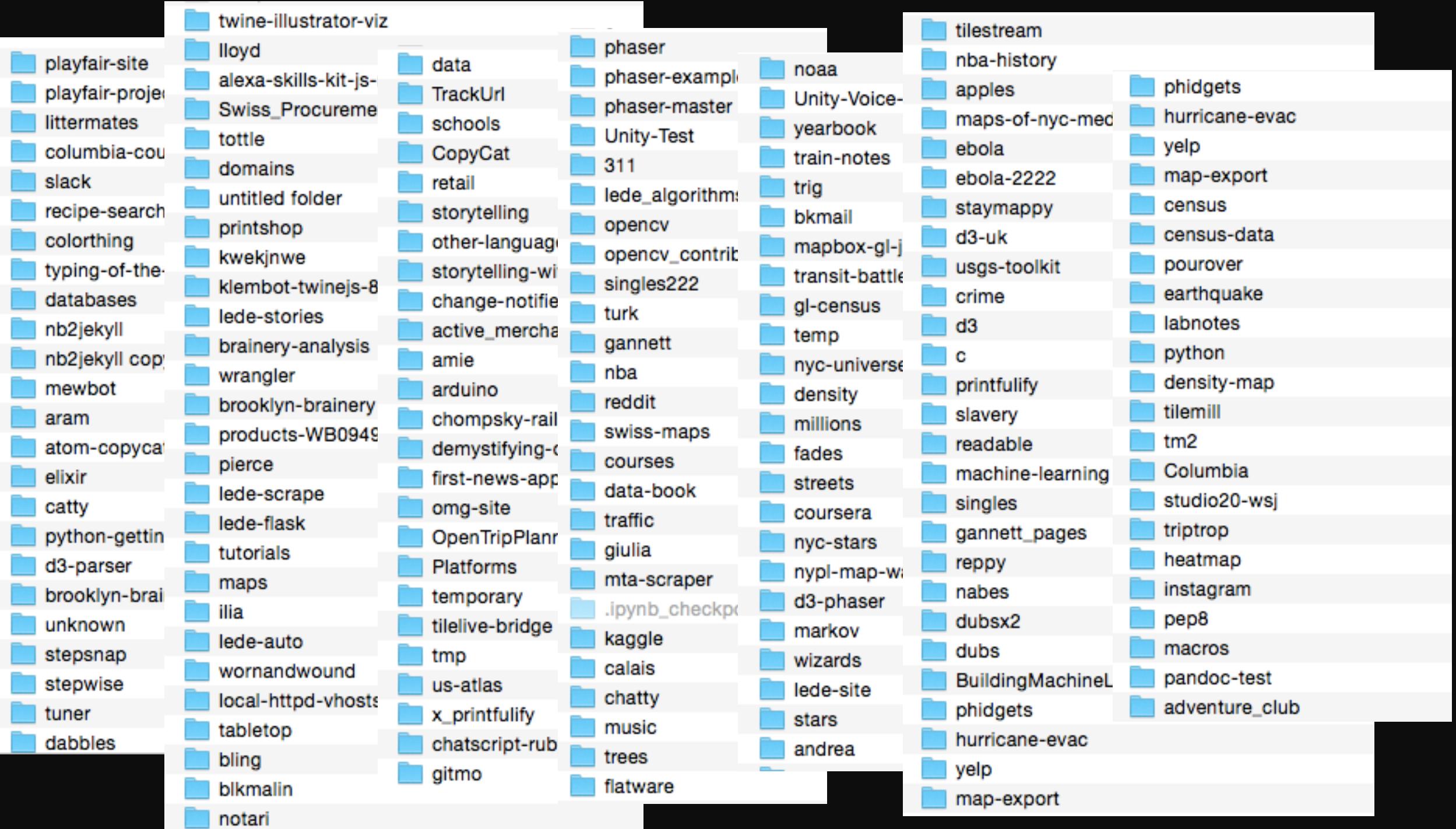
Thinking about mapping life expectancy too, but I'll give it up if it's too much.



JUST GIVE UP

KEEP THINGS MOVING





“ABANDONED” IN STORY.MD

title: How to eat a blueberry properly

summary: Some poeple don't know how to eat a blueberry, so this is a series of visualizations about how to do so

status: abandoned

CRITIQUES

Project 3 first drafts

Beyond visuals, think about alternate data sources, “compared to what?”, options for going deeper/more specific, building on existing work, and **who you might interview about your data/story**

Pick one to discuss, post it in **#data-studio-critique**

INTERVIEWS

What questions do we want to ask?

TAKE A BREAK.

DEBUGGING

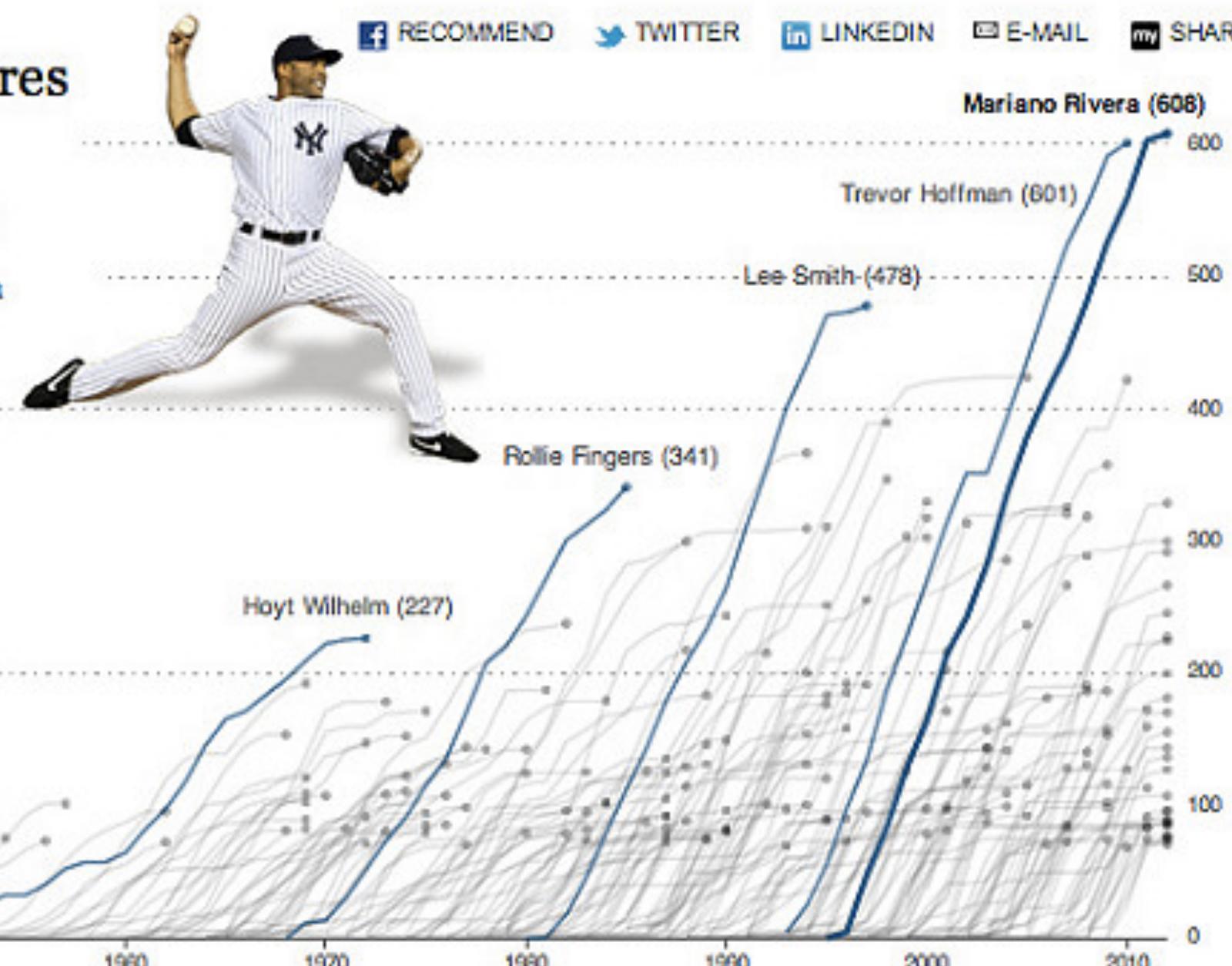
- Shotgun debugging
- Print debugging
- Debugging with pdb
- Using try / except to not debug
- Rubber duck debugging



ANNOTATIONS IN MATPLOTLIB

How Mariano Rivera Compares to Baseball's Best Closers

Mariano Rivera tore his right A.C.L. while shagging fly balls on Thursday, possibly ending his career. Considered the best closer in baseball history, Rivera has more saves than any other pitcher. Below, the cumulative saves of the pitchers with 100 or more.



The closers who broke new hundred-save milestones:

Firpo Marberry (101)
The first reliever to get to 100 cumulative

Hoyt Wilhelm (227)
In addition to being the first pitcher to break the

Rollie Fingers (341)
Known for his handlebar moustache, Fingers was

Lee Smith (478)
From 1983 to 1995, Smith averaged 35

Trevor Hoffman (601)
Hoffman was the first to break the 500 and 600

